The problem of diagnosis with respect to physical parameters for changes in structures

George V. Moustakides

George V. MOUSTAKIDES

THE PROBLEM OF DIAGNOSIS WITH RESPECT TO PHYSICAL PARAMETERS FOR CHANGES IN STRUCTURES

Publication interne
n° 295

Avril 1986

PI 286 **La tolérance aux fautes dans un système temps-réel à contraintes strictes**
Maryline Silly – 32 pages ; Février 86.

PI 287 **A new statistical approach for the automatic segmentation of continuous speech signals**
Régine André-Obrecht – 38 pages ; Mars 86.

PI 288 **Synthèse sur les réseaux locaux temps-réel**
Philippe Belmans – 40 pages ; Mars 86.

PI 289 **Calcul distribué d'un extrémum et du routage associé dans un réseau quelconque**
Jean-Michel Hélary, Aomar Maddi, Michel Raynal – 36 pages ; Mars 86.

PI 290 **A new matrix multiplication systolic array**
Patrice Quinton, Brigitte Joinnault, Pierrick Gachet – 12 pages ; Avril 86.

PI 291 **Une introduction à quelques techniques du contrôle distribué à travers un exemple**
Noël Plouzeau, Michel Raynal, Jean-Pierre Verjus – 22 pages ; Avril 86.

PI 292 **Distributed Synchronization of Parallel Programs : why and how ?**
Patrice Quinton, Jean-Pierre Verjus – 16 pages ; Avril 86.

PI 293 **Sur l'utilisation des séquences multi-images en robotique**
Lionel Marcé – 16 pages ; Avril 86.

PI 294 **Stabilité de l'analyse factorielle des correspondances multiples en cas de données manquantes et de modalités à faibles effectifs**
Habib Bénali – 16 pages ; Avril 86.

PI 295 **The problem of diagnosis with respect to physical parameters for changes in structures**
Georges V. Moustakides – 14 pages ; Avril 86.

IRISA

INSTITUT DE RECHERCHE EN INFORMATIQUE
ET SYSTÈMES ALÉATOIRES

Campus Universitaire de Beaulieu
35042 - RENNES CÉDEX
FRANCE
Téléphone : 99 36 20 00
Télex : UNIRISA 950 473 F
Télécopie : 99 38 38 32

Publication Interne n° 295 - Avril 86 - 14 pages

THE PROBLEM OF DIAGNOSIS WITH RESPECT TO PHYSICAL PARAMETERS FOR CHANGES IN STRUCTURES.

George V. Moustakides

IRISA-INRIA

ABSTRACT

Assuming that a structure is modeled by a tripple M,K,C (mass elasticity and damping matrices), once a detection of a change is signaled, we try to answer the diagnosis problem, i.e. what is the most probable variation in the physical parameters that caused the alarm. The diagnosis method is based on the Instrumental Statistics Detection Method. It basically consists in directing this statistics into certain precomputed subspaces that have a known correspondence with the physical parameters. The subspace that maximizes the statistics is considered as the most probable cause of the signaled change.

RESUME

En supposant qu' une structure est modelisee par un triplet M,K,C (matrices de mass, raideur et amortissement), une fois que une detection de rupture est signalee, on essaye de repondre au probleme de diagnostic , i.e. qui est la variation la plus probable dans les parametres physiques qui est la cause de l' alarme. La methode de diagnostic est basee sur la Methode de Detection des Variables Instumentales. Elle consiste a diriger la statistique dans des sous-espaces precalcules qui ont une correspondance directe avec les parametres physiques. Le sous-espace qui maximise la statistique est considere comme la cause la plus probable du changement signale.

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE
(L. A. 227)
UNIVERSITÉ DE RENNES 1 I. N. S. A. DE RENNES

INSTITUT NATIONAL DE RECHERCHE
EN INFORMATIQUE ET EN AUTOMATIQUE
(LABORATOIRE DE RENNES)

I. INTRODUCTION.

Let us assume that a structure is discribed by the following differential equation

$$\begin{aligned} M\ddot{X} + C\dot{X} + KX &= E_t \\ Y &= LX \end{aligned} \tag{1}$$

where M,K,C are the mass,elasticity and damping matrices, X is the state vector, E_t the input and Y the output (observation). For the three matrices M,K,C we will assume that they are symmetric but the mass matrix is not necessarily diagonal. A linear model of the form of (1) is very common for modeling structures even for cases where partial differential equations are needed. It can be for example the model generated by the Finite Element Method (FEM). Usually the matrices M,K,C generated by FEM have a large dimension (for practical problems). If we try to estimate the natural frequencies using the FEM model, the estimation is good only for the first two or three modes. On the other hand the M,K,C representation has a very direct correspondence with the physical structure, we can easily identify the structure by using the elements of these matrices. For this last reason the M,K,C representation is tractable for the diagnosis problem.

If a large structure is vibrating under excitation and we measure these vibrations we can very seldomly correctly identify by any method more than a very limited number of modes. This is because higher modes are not excited enough and are thus lost in the measurement noise. Thus an M,K,C model even though tractable for a diagnosis, it is too complicated for a detection of changes in vibrations, where a more accurate but reduced representation of the relevant modes is needed. For such a case a reduced AR model is more efficient. It has the advantage that it can be estimated using classical estimation methods

directly on output measurements, something that is not possible for the M,K,C model, where the costly FEM is used. In the next section we will talk briefly about the identification problem and try to relate the AR model with the M,K,C model.

II. THE IDENTIFICATION OF A REDUCED MODEL.

Let us assume that we have available a collection of samples $(Y_1,\ldots,Y_T)$ of the output Y of a system under nominal conditions and also a nominal model of the form of (1),i.e. the matrices M,K,C. What we would like to do is to estimate an AR model, that is a set of matrices $A = [A_m,\ldots,A_1]$ such that V_n defined by

$$V_n = Y_n - A_1 Y_{n-1} - \cdots - A_m Y_{n-m} \tag{2}$$

behaves as a moving average sequence. The AR model defined in (2) has an order which is much smaller than the order of the model in (1), because it contains only the excited modes. Once we have selected an order, there exist several methods for estimating A using the output samples, see for example [1,2]. Let us call A_o the AR model estimated by using the output samples. This model identifies correctly the excited modes and is usually sufficient to discribe the STATISTICAL behavior of the signal output Y. This is the reason why it is used for detection problems.

Let us now consider the estimation of A by using the model in (1). It is clear that the sampled version of the output Y of (1) can be represented as an ARMA process. The problem is that the AR-part will be too large since it contains all the modes of (1) and not only those that are excited. We thus need a model reduction. Let us assume that we have made a selection of eigenvalues and their corresponding eigenvectors. They are usually in complex conjugate pairs. If $(\lambda_1,\ldots,\lambda_{2N})$ and $(\varphi_1,\ldots,\varphi_{2N})$ are the selected eigenvalues and their

corresponding eigenvectors, we will assume throughout this paper that all the selected eigenvalues are simple. For a pair λ_i, φ_i we have

$$(\lambda_i^2 M + \lambda_i C + K)\varphi_i = 0 \tag{3}$$

Also because of the symmetry of the matrices we have

$$\varphi_i^T(\lambda_i^2 M + \lambda_i C + K) = 0 \tag{4}$$

for $i = 1,2,\ldots,2N$. It is now easy to see that if these are the only excited modes then the reduced AR model satisfies

$$(\mu_i^m I - \mu_i^{m-1} A_1 - \cdots - A_m) L\varphi_i = 0 \tag{5}$$

where $\mu_i = exp(\lambda_i \tau)$ and τ is the sampling period. If d is the length of the output vector Y then $d \times m$ must be equal to $2N$; if this is not possible we take m larger in order $d \times m$ just exceeds $2N$. The system of equations defined in (5) can be put under a more compact form

$$LTe^{m\tau D} - A_1 LTe^{(m-1)\tau D} - \cdots - A_m LT = 0 \tag{6}$$

$$\text{where } T = [\Phi \; \Phi'] \quad \text{and} \quad D = \begin{bmatrix} \Delta & 0 \\ 0 & \Delta' \end{bmatrix}$$

Δ is a diagonal matrix containing only half of the selected eigenvalues and Δ' contains the conjugate eigenvalues. Φ is a matrix that contains half of the selected eigenvectors and Φ' the conjugate eigenvectors. Clearly the eigenvectors in Φ correspond to the eigenvalues in Δ. If we now call $B_k = LTe^{k\tau D}$ then (6) can be written

$$B_m^T = B\, A^T \tag{7}$$

where $B = [B_{m-1} \cdots B_0]^T$. We call the matrix A identified using Equation (7),

A_*. Clearly A_0 and A_* are not the same, because as we said in the introduction the FEM is not very accurate for the identification of the modes and also because the identification of A_0 is usually biased because of the model reduction. The only thing one should take care of in order to reduce somehow the difference, is to select those eigenvalues from the M,K,C model that have frequencies close to the ones identified by using the A_0 model. In any case the difference between A_0 and A_* is not very crucial basically because A_* is not used for any test. As we will see, it is used only for computing the different subspaces that are necessary for the diagnosis. In the next section we present the Instrumental Statistics Detection Method (ISDM) and how it can be used to make a diagnosis.

III. THE INSTRUMENTAL STATISTICS DETECTION METHOD.

We briefly present this detection method by insisting more in the property of applying it on a certain subspace. More details regarding theoretical questions and also practical implimentation of the test can be found in [3,4,5]. Its optimality properties with maximization of its power by optimaly positioning the sensors on the state vector can be found in [6,7].

Let us assume that we are given a nominal AR model A_0 and a collection of covariances $R_0,R_1,\ldots$ where $R_k = \sum_{n=1}^{T} Y_n Y_{n-k}^T$. Again $\{Y_n\}$ is the sampled observation sequence. What we would like to do is to test if the observations are compatible with the nominal model A_0. We define the following statistics

$$U_N = H_{m,N-1}^T \begin{bmatrix} A_0^T \\ -I \end{bmatrix} \tag{8}$$

where $H_{p,N-1}$ is the Hankel matrix defined by

$$H_{p,N-1} = \begin{bmatrix} R_0 & \cdots & R_{N-1} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ R_p & \cdots & R_{N+m-1} \end{bmatrix} \tag{9}$$

and N is the number of instruments. The problem of detecting a change, is now equivalent to detecting a change in the AR model A_0. If we follow a local approach, that is if we suppose that the possible changes are of the form $\Theta/\sqrt{T}$ where Θ is a "direction" of change of A_0 then as the number T of observations goes to infinity we can show the following central limit theorem [3].

Theorem. If U_N is defined as in (8) then if H_i, $i=0,1$ denotes the no change and the change hypothesis we have

$$\begin{aligned} H_0&: \quad 1/\sqrt{T}\, U_N \to N(0,\Sigma_T) \\ H_1&: \quad 1/\sqrt{T}\, U_N - 1/T\, H^T_{m-1,N-1}\Theta \to N(0,\Sigma_T) \end{aligned} \tag{10}$$

where the convergence in (10) is in distribution. Thus under H_0 we have that U_N behaves as a centered Gaussian matrix and under H_1 as non centered. The covariance matrix Σ_T can be estimated using the covariances R_k, for details see [4,5]. Clearly now in order to test between the two hypotheses we use (10) as equality and thus the detection problem reduces to detecting a change in the mean of a Gaussian random vector. This problem is a classical detection problem. If we do not know anything about the direction Θ then we maximize the likelihood with respect to Θ and this yields a χ^2 statistics which for the case of $N = m$ (number of instruments equal to the AR order) takes the form

$$T_T = u^T \Sigma_T^{-1} u \tag{11}$$

where u is the vector which we get if we put the columns of U_N one under the other. A change is signaled if T_T exceeds a threshold. The test defined this way is used as "global" since the only information that it gives is

whether there is a change or not. If we like now to test whether the mean has changed according to some prespecified direction Θ with an unknown magnitude δ then we maximize the likelihood only with respect to the scalor δ. This yields the following test statistic

$$T_\Theta = \frac{(u^T \Sigma_{\bar{T}}^{-1} H_{m-1,N-1} \vartheta)^2}{\vartheta^T H_{m-1,N-1}^T \Sigma_{\bar{T}}^{-1} H_{m-1,N-1} \vartheta} \tag{12}$$

where again ϑ is the vector version of Θ. Comparing T_T and T_Θ we can see that

$$T_\Theta \leq T_T \tag{13}$$

The direction Θ does not necessarily correspond to the actual direction of change. In any case, we expect that if the change was indeed in the direction of Θ or to one "very close" to Θ, to have a T_Θ very close to the global test statistic T_T. On the other hand if Θ is "far" from the direction of the actual change to have a T_Θ much smaller than T_T. This idea can be used to test the change of any other parameter related to our system. Assuming that the change of this parameter is small it will produce a variation in the AR model A which will be equal to some matrix Θ times an unknown magnitude. This matrix Θ is the derivative of the AR model with respect to the parameter under consideration (first order approximation of A around A_0). This diagnosis method is called "sensitivity method" for the obvious reason and was applied in [4,5] to test which modes have been changed, once an alarm is signaled. We will use the same idea for detecting changes in the M,K matrices.

IV. THE SENSITIVITY METHOD FOR THE M,K MATRICES.

As we have seen in section III all we can detect is changes in the AR model. So any change we like to test must be transformed into a corresponding change

in the AR parameters. In other words if we like to test if we have a specific change in a mass or an elasticity we have to compute the derivative Θ of the AR model A with respect to the corresponding mass or elasticity. Unfortunately such a computation needs not only the knowledge of the the excited eigenvalues and the observable part of their corresponding eigenvectors (things that can be computed directly from A_0 and thus from the data); but rather the knowledge of the whole eigenvectors. This is where the model M,K,C and A_* comes in. In other words we will not compute the derivative of A_0 but the derivative of A_* defined by equation (7). Taking thus derivatives in (7) gives

$$(\delta B_m^T) = (\delta B)A_*^T + B(\delta A_*^T) \tag{14}$$

where δX denotes the derivative of X with respect to the quantity we want. In (14) δA_* is exactly the direction we like to compute (what we call Θ). From the definition of B we see that it is enough to compute δB_k. Since $B_k = LTe^{k\tau D}$ taking derivatives gives

$$\delta B_k = L(\delta T)e^{k\tau D} + k\tau e^{k\tau D}(\delta D) \tag{15}$$

From (15) we see that in order to compute the derivative of B_k we need the derivatives of the eigenvalues and the eigenvectors. In order to make this computation we use a method similar to the one used in [8].

Computation of changes of eigenvalues and eigenvectors.

Let λ and φ be an eigenvalue and its corresponding eigenvector then we have that

$$(\lambda^2 M + \lambda C + K)\,\varphi = 0 \tag{16}$$

Taking derivatives in (16) gives

$$(\delta\lambda)\,(2\lambda M + C)\,\varphi + (\lambda^2 M + \lambda C + K)\,(\delta\varphi) + (\lambda^2(\delta M) + (\delta K))\,\varphi = 0 \qquad (17)$$

In (17) we have finally the quantities δM and δK we are interested in. Now multiplying (17) from the left by φ^T and using (4) gives

$$\delta\lambda = \frac{\varphi^T\,(\lambda^2\,(\delta M) + (\delta K))\,\varphi}{\varphi^T\,(2\lambda M + C)\,\varphi} \qquad (18)$$

Using the expression of $\delta\lambda$ from (18) in (17), gives a linear system in $\delta\varphi$. This system has not a unique solution because if $\delta\varphi$ is a solution so is $\delta\varphi + \alpha\varphi$. We can thus look for the solution that is orthogonal to φ, i.e.

$$\varphi^T\,\delta\varphi = 0 \qquad (19)$$

We can now substitute for example one of the equations in (17) with (19) and thus getting a system with a solution that is orthogonal to the eigenvector φ. Actually condition (19) is equivalent to saying that we are looking for that changed eigenvector $\varphi + \delta\varphi$ that is as close as possible in the euclidean sense to φ. In any case we can verify that by taking any solution of (17) does not change δA_* computed from (14).

Using Equations (17,18,19) we can find δD and δT and then from (14) δA_* for different forms of δM and δK matrices. The most common form for δM is $e_i\,e_i^T$ for a change in the i-th diagonal element (e_i is a vector with all its elements equal to zero except the i-th which is equal to unity). For δK we have $(e_i - e_j)\,(e_i - e_j)^T$ for a change in the cross elasticity between points i and j, or $e_i e_i^T$ for the auto elasticity at the point i.

Let us now assume that we have a collection of directions $\Theta_1, \ldots, \Theta_L$ corresponding each one to some change in the matrices M,K. Since in general the representation M,K,C has much more parameters than the representation of A we expect that some of the directions Θ_i will fall close to each other and

our sensitivity test statistic T_Θ defined by (12) we will not be able to distinguish between them. We thus need to classify our directions into clusters and for each cluster to use a representative. In defining clusters we need first to define a distance measure that has some meaning with respect to our test.

Definition of a Distance Measure.

As we said the measure we want to define must have a close relation with our diagnosis method. Let us thus assume that we have an actual change in the direction Θ_i and that we check in the direction Θ_j using the statistics defined by (12). First notice that in deriving the test we assumed that the change is of the form $\Theta/\sqrt{T}$. That is, the magnitude of the change goes to zero as the number of observations increaces. This assumption is always made whenever we follow a local approach. In practice though it is nonrealistic since a change has usually a given magnitude. Assuming that this is the case, the random matrix U_N will have a mean that will go to infinity as the number of observations increases and thus this matrix will fluctuate around this non random value that goes to infinity. Since the variance is finite, in the limit U_N properly normalized by T will tend to its mean. Thus in the expression of the statistics (11) and (12) we can substitute U_N with its mean. This gives

$$T_T \sim \vartheta_i^T H^T \Sigma^{-1} H \vartheta_i$$

$$T_{\Theta_j} \sim \frac{(\vartheta_i^T H^T \Sigma^{-1} H \vartheta_j)^2}{\vartheta_j^T H^T \Sigma^{-1} H \vartheta_j} \tag{20}$$

Considering in (20) the approximate equalities as equalities and computing H and Σ using the observations under H_0 we have again $T_\Theta \leq T_T$. We can define now the quantity

$$\rho_{i,j} = \frac{(\vartheta_i^T H^T \Sigma^{-1} H \vartheta_j)^2}{(\vartheta_i^T H^T \Sigma^{-1} H \vartheta_i)(\vartheta_j^T H^T \Sigma^{-1} H \vartheta_j)} \tag{21}$$

as a proximity measure between Θ_i and Θ_j. A distance measure can be easily defined from $\rho_{i,j}$. For example we can use $d_{i,j} = 1 - \rho_{i,j}$. Using this definition of distance we can use classical clustering algorithms and classify our directions. Then to each class we correspond a representative (the mean direction for example) and use these representatives to make a diagnosis. Clearly by selecting a specific representative as the cause of the alarm we select a whole class of possible changes whithout being able to distinguish among them. For certain applications this is enough as information.

V. CONCLUSION.

We have presented here a method of diagnosis with respect to the physical parameters of a structure when an alarm for a change is signaled. This method is very closely related to the Instrumental Variable Detection Method and actually directs the test statistics in certain precomputed subspaces. The relation between the subspaces and the physical parameters is known apriori, so the selection of a specific subspace corresponds to a selection of a set of possible physical parameters. As it is seen in Section IV, we are interested only in changes with respect to the mass and elasticity matrix elements. This is not restrictive since for large structures the damping matrix is very close to zero and considering changes with respect to this matrix is useless.

REFERENCES

[1] M. PREVOSTO, "Algorithmes d' Identification des Characteristiques vibratoires de structures mecaniques complexes", *These pour le titre de Docteur Ingenieur*, Universite de Rennes I, Mars 1982.

[2] J-J. FUCHS and A. BENVENISTE, "Single Sample Modal Identification of a Nonstationary Stochastic Process", *IEEE Trans. on Automatic Control*, Vol. AC-30, No 1, pp. 66-74, Jan. 1985.

[3] G.V. MOUSTAKIDES and A. BENVENISTE, "Detecting Changes AR Parameters of a Nonstationary ARMA Process", *Stochastics*, Vol. 16, pp. 137-155, 1986.

[4] M. BASSEVILLE, A. BENVENISTE, G.V. MOUSTAKIDES, "Detection and Diagnosis of Abrupt Changes in Modal Characteristics of Nonstationary Digital Signals", *IEEE Trans. on Inf. Theory*, to appear in May 1986.

[5] M. BASSEVILLE, A. BENVENISTE, G.V. MOUSTAKIDES, A. ROUGEE, "Detection and Diagnosis of Changes in Eigenstructure of Nonstationary Multivariate Systems", *IRISA, Publ. Interne*, No 276, Dec. 1985.

[6] A. ROUGEE, M. BASSEVILLE, A. BENVENISTE, G.V. MOUSTAKIDES, "Optimum Robust Detection of Changes in AR Part of a Multivariate ARMA Process", *IRISA, Publ. Interne*, No 277, Dec. 1985.

[7] M. BASSEVILLE, A. BENVENISTE, G.V. MOUSTAKIDES, A. ROUGEE, "Optimal Sensor Location for Detecting Changes in Dynamical Behavior", *IRISA, Publ. Interne*, No 285, Feb. 1986.

[8] STEWART, "*Introduction to Matrix Computation*", Acad. Press, 1973, pp. 289-306.